Activities... what did we learn?

larger n --> * less variability (smaller spread)
* more normal shape

all sample sizes, all activities -->

* average (% or mean) was close to the true value (% or mean)

WARM UP:

The weights of Labradors follow a normal model with a mean of 42 lbs. and std. dev of 3.4 lbs.

$N(42, 3.4)$

1. What is the probability that you have a dog that weights more than 50 lbs? $P(X>50) = 0.009$
2. What is the probability that you have a dog that weights less than 53 lbs? $P(X<53) = 0.994$
3. What is the probability that you have a dog that weighs between 41 and 48 lbs?

$P(41 < X < 48) = 0.576$

Ch. 18: Central Limit Theorem

Sampling Distribution-

* A histogram of repeated samplings
* We can make histograms with means or proportions (%)
* Examples: true % color 1 in foam pieces
  true avg. score on SATs
* We use these histograms to help us find the TRUE mean (μ) or proportion (p)

CENTRAL LIMIT THEOREM: (CLT)

* A *sampling distribution* can be approximated by a Normal model

* The larger the sample, the better the approximation will be n=5 n=20

* Does not matter what shape the population is... if the sample size (n) is large enough, the sampling distribution --> Normally shaped

Ex: pennies

* EXAMPLE: BINOMIAL

PROPORTIONS:

A sample of size n is taken from a population with center *p (true prop. or %).* $p = 0.25$

Sampling distribution of sample proportions ($\hat{p}$):

Mean: $\mu_{\hat{p}} = p$ 0.25 Std. Dev: $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$ $\sqrt{\frac{(0.25)(0.75)}{20}}$

*** On formula sheet*

*According to the CLT, if *n* is large enough then the sampling distrib. is approx. normal.

*Also, the following conditions must be met: (STATED&CHECKED)

State	Check
1) Simple Random Sample (SRS)	1) Stated SRS Assumed representative
2) np and nq ≥ 10	2) $(100)(0.30) \geq 10$, $(100)(0.70)$
3) population ≥ 10*n*	3) there are more than 1000 people

** If all these are met, we can say our sampling distribution is NORMAL -->*

$$N\left(p, \sqrt{\frac{pq}{n}}\right)$$

Example: According to the manufacturer of the candy Skittles, 20% of the candy produced is the color red. What is the probability that given a large bag of skittles with 58 candies that we get at least 17 red (0.293 red)?

$p = 0.20$

$n = 58$

$\hat{p} = 17/58 = 0.293$

$P(\hat{p} \geq 0.293) = ?$

State

1) SRS

2) $np \geq 10$, nq

3) $pop \geq 10n$

Check

1) Assumed large bag is representative of all skittles.

2) $(58)(0.20) \geq 10$, $(58)(0.80)$

3) There are more than 580 skittles.

Conditions met, Normal model, $N\left(0.20, \sqrt{\frac{(0.2)(0.8)}{58}}\right)$

$P(\hat{p} > 0.293) = 0.0383$

normalcdf$\left(0.293, E99, 0.20, \sqrt{\frac{(0.2)(0.8)}{58}}\right)$

MEANS:

A sample of size n is taken from a population with center μ and std. dev. of σ.

Sampling distribution of sample means ($\bar{x}$):

Mean: $\mu_{\bar{x}} = \mu$ Std. Dev: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$\mu = 2.25$ $\sigma = 1.393$

$\sigma_{\bar{x}} = \frac{1.393}{\sqrt{5}} = 0.623$

* According to the CLT, if n is large enough then the sampling distrib. is approx. normal.

* Also, the following conditions must be met: (stated&checked)

1) Simple Random Sample (SRS)

2) $n \geq 30$ or the **population** must be normally distributed

3) population $\geq 10n$

*** If all these conditions are met, we can say that the sampling distribution is NORMAL -->*

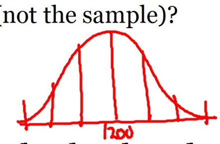
$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Example: Suppose that male seniors have a mean SAT score of 1200 with a standard deviation of 130. We take a random sample of 100 male seniors.

$\mu = 1200$ $\sigma = 130$ $n = 100$

(a) If we assume a normal population, what would the distribution be for the male seniors (not the sample)?

$N(1200, 130)$



(b) What is the probability that **a randomly selected senior** scores less than 1150?

$P(X < 1150) =$ normalcdf$(-E99, 1150, 1200, 130)$

0.3502

(c) What would the distribution of sample means look like?

State

1) SRS

2) $n \geq 30$ or normal pop.

3) $pop \geq 10n$

Check

1) stated random sample

2) $n = 100 \geq 30$ stated normal pop.

3) there are more than 1000 male srs.

$\sigma/\sqrt{n}$

Conditions met, Normal model, $N\left(1200, \frac{130}{\sqrt{100}}\right)$

(d) What would be the probability that we would get an **average** for the sample of male seniors of less than 1150?

$P(\bar{X} < 1150) =$ normalcdf$\left(-E99, 1150, 1200, \frac{130}{\sqrt{100}}\right)$

$= 6.002 \times 10^{-5}$

p. 434 #15, 37, ~~39~~

1) Write important info (n, p, std. dev, x-bar, etc.)
2) STATE and CHECK conditions
3) If conditions pass, write "conditions met" and then write the model you are using: N(,)
4) Solve problem (use prob. notation!)

15) n = 200 p = 0.07

(b) Conditions:

(1) SRS	(1) assumed representative
(2) np & nq ≥ 10	(2) (200)(0.07) ≥ 10 (200)(0.93)
(3) pop ≥ 10n	(3) all loans by this bank ≥ 2000

Conditions met => $N\left(0.07, \sqrt{\frac{(0.07)(0.93)}{200}}\right) = N(0.07, 0.018)$

(a) Mean and std. deviation given in model in (b)

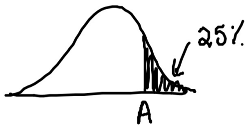
(c) $P(\hat{p} > 0.10)$ = normalcdf(0.10, E99, 0.07, 0.018) = 0.0482

37) N(266, 16)

(a) $P(270 < X < 280)$ = normalcdf(270, 280, 266, 16) = 0.2105

(b) $P(X > A) = 0.25$ invnorm(0.75, 266, 16)

A = 276.79 days



(c) n = 60

Conditions:

(1) SRS	(1) assumed representative
(2) n ≥ 30 OR pop normal	(2) n = 60 ≥ 30 stated normal pop
(3) pop ≥ 10n	(3) ~~all women cared for by this obstetrician ≥ 600~~ there are more than 600 preg. women

conditions met => $N\left(266, \frac{16}{\sqrt{60}}\right) = N(266, 2.0656)$

(d) $P(\bar{x} < 260)$ = normalcdf(-E99, 260, 266, 2.0656) = 0.0018

HW: p. 434
#16, 20, 22, 38, 48